THE GERMPLASM RESOURCES INFORMATION NETWORK - GRIN

A BRIEF HISTORY AND INTRODUCTION

UNITED STATES DEPARTMENT OF AGRICULTURE

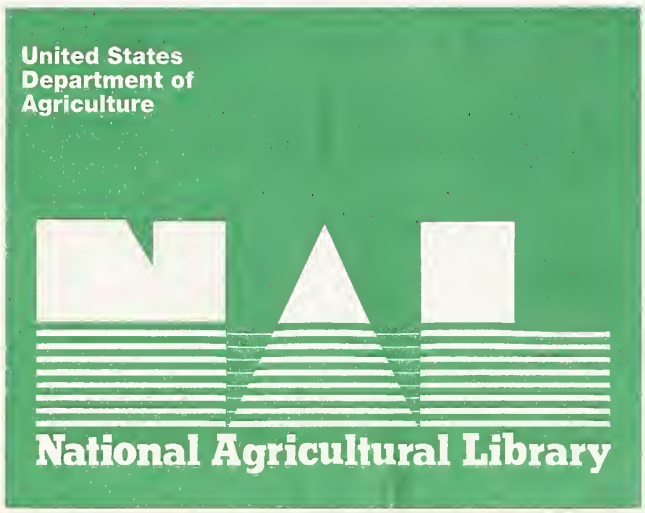
AGRICULTURAL RESEARCH SERVICE

PLANT GENETICS AND GERMPLASM INSTITUTE

DATABASE MANAGEMENT UNIT

---

31 JANUARY 1987


United States
Department of
Agriculture
National Agricultural Library

WHAT IS GRIN?

Plant genetic resources within the United States are managed by the National Plant Germplasm System (NPGS). One of the major structural components of this system is the Germplasm Resources Information Network (GRIN). GRIN is a centralized computer database used to facilitate management and operation of the NPGS and to enhance communication to scientists regarding the location and characteristics of material they may wish to obtain for research purposes. The major purpose of the GRIN is to serve as a central repository of information concerning both major and minor aspects of plant genetic resources within the NPGS and to provide ready accessibility of this information to all users of the system.

WHY HAVE GRIN?

Importance of Genetic Resources

Plant genetic resources (or germplasm) is the raw material required by plant breeders for the development of new, superior crop varieties that can ensure a stable, plentiful supply of high quality food, feed, and fiber. Most of the plants on which the United States depends were introduced from other countries and developed to suit the environment in which they would be grown. The list of economically important native plants is very short and includes sunflowers, cranberries, blueberries, strawberries, pecans, hops, range grasses, conifers, and hardwoods.

There are large gaps in the base of genetic diversity of some crops, particularly the wild species and primitive varieties. These rich sources of variation may contain genes for disease and insect resistance and other desirable traits, but in many areas of the world, these sources of diversity are rapidly being depleted, displaced, or abandoned. Once lost, these sources will never again be available to mankind. The need for this diversity becomes apparent when the genetic vulnerability of present American monoculture is measured against the constant battle against plant pathogens and pests.

Plant Introductions and the NPGS

The American Government recognized the need for a continuing search for more adaptable crops early in its history. In 1891, American overseas consuls were urged to send useful plants back to the United States. From this start, the essential elements of the present plant germplasm system gradually developed. This system evolved into the NPGS. The goal of the NPGS is to provide, on a continuing, long term basis, the plant genetic diversity needed by farmers and public and private plant scientists to improve productivity of crops and minimize their vulnerability to biological and environmental stresses.

Minimizing crop losses through control of major stresses is far more difficult and costly than increasing the genetic diversity among varieties of a given crop. Therefore, an NPGS objective is to broaden the genetic diversity of a crop throughout its production areas by having that production come from an array of varieties, all productive, but each different from the others in its range of tolerance to one or more potential stresses. Collection and introduction of new germplasm is the first step toward achievement of this goal.

The NPGS now maintains over 400,000 accessions (samples) of germplasm in the form of seed and vegetatively propagated stocks. These accessions are primarily landraces and unimproved material from foreign sources. New accessions are added to the NPGS at a rate of 7,000 to 15,000 per year. The need for the actual accession to enter the germplasm system is paralleled by the need for information about the accession to be available to users of the system. The immense size of this system creates a challenge for information management. Difficulty of obtaining information, lack of uniformity concerning this information, and its overall poor treatment prompted the NPGS to integrate an information management system as a major structural component. GRIN was designed and developed to act as this system.

BRIEF HISTORY OF GRIN

A feasibility study was conducted during 1976-77 which investigated and identified the need for information management systems that enhanced information availability with regard to collection, conservation, d distribution, and utilization of plant genetic resources within the NPGS. This study drew the following conclusions about the existing information management system: "An information system exists within the plant genetic resources community of the United States but this system lacks the organization, communication techniques, trained personnel, and funding to satisfy the requirement of the NPGS community." From this study the USDA Agricultural Research Service (ARS) recognized the critical need for a nationally unified information system to serve the diverse needs of the NPGS. GRIN was the result of this realization. The developmental phase of this system was the Germplasm Resources Information Project (GRIP). This project was established under a 5-year cooperative agreement with the Laboratory for Information Science in Agriculture (LISA) to develop a computer based information system.

Analysis of the diverse needs of the germplasm community identified two groups of information users within the NPGS. The suppliers consist of those who acquire, maintain, and distribute genetic resources and data. It includes curators and staff of the National Seed Storage Laboratory (NSSL) and Regional Plant Introduction Stations (RPIS). A second, or demand group, consists of plant breeders, scientists, and other researchers from both the public and private sectors who use the genetic resources and data.

Further analysis identified specific needs of both groups. Small-scale prototypes were then constructed to meet the needs of each group on the supply side as well as to verify their objectives. From the evaluation of these prototypes, a user oriented approach was

selected for design development.

In the evaluation and design of the information system, a centralized computer (Appendix 1) was selected to optimize operational speed for all users. A database concept was adopted to enhance information by reducing redundancy and by relating most of the pertinent information about a particular accession--from its native habitat to the most recent characteristic and evaluation results. This centralization also allows researchers access to a more extensive collection of samples from which to choose. This reduces the possibility of overlooking a potentially valuable sample. Maintenance of information is supported through updates that are quickly available to everyone. The most accurate and current information is thus accessible without time-consuming written notifications. Information updates are made by individuals and organizations recognized as experts in their particular area(s). For instance, plant taxonomists monitor and maintain taxonomic nomenclature; the Plant Introduction Office (PIO) in Beltsville, Maryland, maintains information concerning origin and particulars about introduction; the RPIS and other germplasm repositories maintain viable samples and serve as points of contact for sample availability and characteristic information; and the breeders, growers, and researchers provide detailed evaluation information.

To make the data understandable and consistent at a national level, Crop Advisory Committees (CAC's) were developed simultaneously with GRIP. These committees, composed of crop experts from public and private sectors of the NPGS, develop evaluation and characterization criteria as well as descriptor lists and standard methods of measurement and reporting.

The completed design phase brought the transformation of GRIP to GRIN. On July 1, 1983, GRIN was transferred to the Plant Genetics and Germplasm Institute (PGGI) within the ARS, USDA. This Institute is located in Beltsville, Maryland. The management and final development of the network are controlled by the GRIN Database Management Unit (DBMU). Implementation of this system (called GRIN1) was finalized in February 1984.

Soon after the implementation phase was finished, it was realized the system design was not sufficient to adequately include all parts of the NPGS and to accommodate the vast array of activities presented by this system. Major shortcomings included: omission of three major functional groups from the original design; inadequate germplasm ordering and inventory procedures; the use of too few data fields for the description of an accession; the lack of fast computer access to heavily used data; and a lack of fexibility and overall efficiency of the database. Another assessment of the needs was performed during 1985, when germplasm collection site personnel from 12 locations were brought to Beltsville to finalize the updated design. Expansion of the GRIN1 design and other programming will be concluded in March 1987. Features of the enhanced database design (GRIN2) compared to GRIN1 are: conversion of PIO, taxonomic support, and NSSL from their autonomous computer systems to GRIN; provision for germination rules and results; a simplification of the adding, modifying, and deleting of germplasm collection sites and associated inventory records; a more flexible characteristic and evaluation data structure; inclusion of germplasm ordering capabilities; a revision of and increase in the number of data fields; a doubling in size of the GRIN1 database; the inclusion

of a revised friendly menu system for collection site users; a provision for collection site users to load their own data into the database; the inclusion of a number of predefined data reports; and an upgrade of security that increased efficiency of the database.

HOW DOES GRIN WORK?

GRIN has three important functions to fulfill. First, it serves as a central repository for valuable genetic resources information that is accessible by the entire scientific community. Second, it is a means for the CACs to begin standardization of crop descriptors and evaluation information. Third, it provides a mechanism for each of the RPIS and other sites to manage daily inventory.

Anyone who can justify a need for accessing the GRIN database can request access by writing the GRIN DBMU. Logon identification and an access code are then assigned by the DBMU. Also, documentation and instructions for use of the system are supplied. Foreign scientists affiliated with any International Agricultural Research Center may also gain access through this procedure.

The database is designed to permit flexibility to the users in storing and retrieving information. A network design is one that allows multiple paths to the data but has linkages that connect all the data together. Figure 1 shows how different information is related. Appendix 2 contains a list of all data fields or elements that compose the GRIN2 database.

The public user system is designed for the inexperienced user and accommodates most hardware types that have telecommunication capabilities. Public users are presented with a series of menus which offer quick, easy access to major GRIN features and database searches. The menus assist unfamiliar users in the basic features of GRIN. Experienced users are also free to exercise GRIN features beyond the limits of the options presented through the use of menus.

The maintenance of data within GRIN is the responsibility of various functional groups within the NPGS. The PIO has sole responsibility for maintaining accurate passport data, geographic acquisition, and geographic origin information. USDA taxonomists maintain all plant taxonomic information. PIO and the taxonomists are able to modify information in their respective areas, however, any user possesses retrieval access. The germplasm collection sites are permitted to modify passport information for any accession with a primary identifier that does not have a "PI" prefix, or any sample with a "PI" prefix and number less than 431464. The RPIS and other participating germplasm collection sites are reponsible for maintaining accurate inventory and characteristic data for their respective collection. Public users are permitted to retrieve and examine all information from the database.

The DBMU acts as the caretaker of GRIN. This responsibility includes the maintenance of: all computer application software (programs); the database management system (DBMS); and a liaison with computer operations (the Prime minicomputer). The majority of volume data loading and the compilation and writing of GRIN documentation also lies with the DBMU. The DBMU also provides technical assistance to

users in preparation of sof'ware that is unique for a specific collection site. Database access and system security are also important system management tasks.

Contact the following office for any additional information:

Database Manager
USDA/ARS/BA/PGGI
BARC-W
Building 001, Room 130
Beltsville, Md. 20705
USA
Phone: 301-344-3318
FTS: 8-344-3318